OFFICE OF NAVAL RESEARCH

GRANT N00014-89-J-1178

R&T CODE 413Q001-05

TECHNICAL REPORT NO. #52

CHARACTERIZATION OF THE Si/SiO_2 INTERFERE MORPHOLOGY FROM QUANTUM OSCILLATIONS IN FOWLER-NORDHEIM TUNNELING CURRENTS

J.C. Poler, K.K. McKay and E.A. Irene
Department of Chemistry
University of North Carolina at Chapel Hill
Chapel Hill, NC 27599-3290

Submitted to:

Journal of Applied Physics

REPORT DOCUMENTATION PAGE

Form Approved
OMB No 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	5/4/93	

4. TITLE AND SUBTITLE
Characterization of the Si/SiO_2 Interfere Morphology From Quantum Oscillations in Fowler-Nordheim Tunneling Currents

5. FUNDING NUMBERS
#N00014-89-J-1178

6. AUTHOR(S)
J.C. Poler, K.K. McKay and E.A. Irene

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
The University of North Carolina
Chemistry Department
CB #3290 Venable Hall
Chapel Hill, NC 27599-3290

8. PERFORMING ORGANIZATION REPORT NUMBER
Technical Report #52

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Office of Naval Research
600 N. Quincy Street
Arlington, VA 22217-5000

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
None

12a. DISTRIBUTION/AVAILABILITY STATEMENT
This document has been approved for public release and sale, distribution of this document is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

As design rules shrink to conform with ULSI device dimensions, gate dielectrics for MOSFET structures are required to be scaled to below ~60A where some properties of the device, such as interface roughness, that are negligible for thicker films become crtiical. Microroughness at the interface of ultrathin MOS capacitors has been shown to degrade these devices.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Quantum Oscillations	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	

Your copy

Characterization of the Si/SiO_2 interface morphology from quantum oscillations in Fowler-Nordheim tunneling currents

J.C. Poler, K.K. McKay and E.A. Irene, Department of Chemistry, Venable Hall, University of North Carolina Chapel Hill, NC 27599-3290

As design rules shrink to conform with ULSI device dimensions, gate dielectrics for MOSFET structures are required to be scaled to below ~60Å where some properties of the device, such as interface roughness, that are negligible for thicker films become critical. Microroughness at the interface of ultrathin MOS capacitors has been shown to degrade these devices.

The present study focuses on the interfacial region of ~50Å SiO_2 on Si, using the quantum oscillations in Fowler-Nordheim tunneling currents as a probe. The oscillations are sensitive to the electron potential and abruptness of the film and interfaces. In particular, inelastic scattering and/or thickness inhomogeneities in the film will reduce the amplitude of the oscillations. We are using the amplitude of the oscillations to examine the degree of microroughness at the interface that results from a pre-oxidation high temperature anneal in an inert ambient containing various amounts of H_2O. AFM imaging has shown correlations supporting a microroughness induced change in the quantum oscillation amplitudes.

I. BACKGROUND

As gate oxide thicknesses in MOSFET devices are decreased to less than 100Å, the nature of the material can no longer be appropriately described by its bulk properties. Chemical transition regions at the Si/SiO_2 interface[1], density gradients due to residual stress[2] and interface morphology can all play a significant role in determining the electrical characteristics of the insulating thin film. This can in turn effect device performance and reliability.

As insulator film thickness decreases, transport due to direct and Fowler-Nordheim (FN) conduction dominates and device characterization via the usual C(V) and I(V) analyses becomes difficult. In the present study we use high field FN tunneling characteristics, specifically quantum oscillations in the FN regime, to extract information about the quality of the $Si-SiO_2$ interface in MOS device.

A comparison of the measured I(V) for two MOS tunneling capacitors with virtually identical 55Å thick SiO_2 films (as determined by ellipsometry assuming a constant film refractive index of $N_1 = 1.465 - 0i$) is shown in Fig. 1. The only significant difference in the processing of the samples is the type of pre-oxidation anneal the Si received before the film was grown. The sample as measured by the upper curve (solid circle) was exposed to a high temperature dry inert gas anneal for 30s before the oxygen was turned on in the growth furnace. Even though this sample has the same optical thickness as the sample that did not receive a dry pre-oxidation anneal, the I(V) characteristic shows excess current and this indicates that there is a second transport mechanism that is not described by direct tunneling

through a 55Å barrier. We believe that the excess current measured in the dry annealed sample is due to either an impurity in the film[3] resulting in a "short circuit" or direct tunneling through a thin region of the film resulting from interface roughness of the substrate[4]. It is also noted that in virtually all cases of dry pre-annealed samples, dielectric breakdown occurred prior to the onset FN tunneling.

By using a buffered hydrofluoric acid (BHF) solution the native oxide can be removed from our Si substrates leaving a Si surface that is passivated by a layer of hydrogen atoms[5,6]. At elevated temperatures the hydrogen desorbs from the Si leaving the surface bare and reactive. The Si-H surface species can exist as a mono-, di- and tri-hydride where the terminating Si atom is bound to one, two and three hydrogens respectively. The hydrogen from the di-hydride species will desorb, in vacuum, at 550-600°C and the monohydride will desorb at 700-750°C[7]. Therefore, as the temperature is raised above these thresholds, the surface will depassivate and consequently be highly reactive. If there are oxygen containing impurities present in the hot ambient gas, then the Si will oxidize. Smith and Ghidini[8] (S&G) showed that there are two competing chemical reactions of Si with O_2. For low temperatures and high partial pressures of O_2 in the reaction furnace, Si and O_2 will combine to form SiO_2:

$$Si(s) + O_2(g) \rightarrow SiO_2(s) \quad . \qquad (1)$$

However, for high oxidation temperatures and low partial pressures of oxidant the dominant oxidized species will be SiO, such that:

$$2Si(s) + O_2(g) \rightarrow 2SiO(g) \quad , \qquad (2)$$

where the monoxide is volatilized at elevated temperatures and desorbs from the surface. If an oxide film grows, its thickness will be self controlled once oxide transport is diffusion limited, and the Si/SiO_2 interface will tend toward planarity. For the monoxide production reaction, there is not an overlayer that the oxidant must traverse. Therefore, the morphology of the surface is more subject to external conditions and impurities. Impurities such as SiC can block the oxidation reaction, or agglomeration of oxidant on the surface can enhance the reaction, and it is likely that neither will contribute uniformly across the wafer. The spatial non-uniformity of the oxidation reaction rate at the surface will contribute to the roughening of the substrate.

A thermally grown SiO_2 film will typically grow coplanar to the underlying substrate. For a relatively flat substrate morphology the film thickness will be uniform throughout a large area of the surface. However, if the surface contains terraces or roughness, the oxidation process will tend to smooth out the interface, and consequently the film thickness will not be uniform. Thermodynamically, a system will tend to minimize its free energy. The surface energy at an interface, G_{surf}, is proportional to the surface area. Therefore, to minimize the surface energy the system will try to remove any roughness and the morphology will tend toward planarity.

For thinner oxide films (e.g. <500Å at 800°C) the oxygen can easily traverse the film, and the growth rate is limited by the interfacial reaction[9]. Since there will be an ample supply of oxidant at the interface, the reaction will occur with equal rates (neglecting crystallographic chemical potential differences) at all positions. Therefore, the reaction front will propagate the interface parallel to the local surface normal. As the oxidation

progresses, the vertical surface components, i.e. the sides of the asperities, will be displaced horizontally. When two planes of horizontal oxidation propagation meet, the reaction will tend to remove the horizontal components of the surface. In general, this model predicts that the film grown above an upward protrusion will be thicker than the average thickness, and oxide grown above a downward protrusion, or etch pit, will be thinner.

Etching of the Si surface due to SiO production and sublimation at low oxidant concentrations will result in a roughened surface[10]. By the oxidation mechanism above, subsequent oxidation of the surface could result in an oxide that is inhomogeneous in thickness. Some authors have shown that a high temperature anneal of an oxide covered Si substrate will reduce irregularities in the surface morphology[11]. The oxidation model presented above is consistent with this type of behavior. Other authors have theorized that the oxidation process will roughen a flat surface, and that the roughness can be described by a Poisson distribution of local oxide thicknesses[12,13]. There is probably a combination of effects occurring and the surface morphology will evolve in a way that depends on the specific processing conditions to which it is exposed.

Depending on the shape, height, width of the barrier, the applied bias and temperature of the sample, the tunneling currents will exhibit different dependencies on the electric field across the device. For thin barriers low temperatures and small applied bias the I(V) characteristic is nearly ohmic and exponentially dependent on the insulator thickness[14]. At higher bias the dependence on voltage is not as simple, and the analytical form for the I(V) characteristic depends on the approximations made when integrating the transmission coefficient over the density of states as a function of energy. Stratton[15] shows

that for an arbitrary potential barrier shape, the I(V) characteristic can be approximated by a function that has a non-trivial proportionality on the applied bias in the pre-exponential term, and is exponentially dependent on the bias and the barrier thickness. Most of these experiments are on MOS capacitors with an insulator thickness of about 50Å, that are measured at or below room temperature and where the applied bias is on the same order of magnitude as the work function potential energies of the contacts. The I(V) characteristic for our samples can be described in two regions of applied bias: direct tunneling currents at low bias, and the (FN) tunneling currents at higher bias.

Direct tunneling through the insulating film occurs when electrons are injected from Fermi level of the cathode through the film into empty electron states in the collecting contact. For the MOS capacitor, direct tunneling is dominant for applied biases that produce an energy shift of the contacts less than the electron work function energy of the oxide/anode contact. Under these conditions, the tunneling distance for the electron is constant and equal to the physical thickness of the oxide film. Numerical calculations of I(V) for applied biases between 2-4V, on the poly-Si/(50Å)SiO_2/Si MOS device, shows a nearly linear dependence of voltage on the logarithm of the tunneling currents.

Direct tunneling currents are used, illustrated in Fig. 1 (closed circle), to analyze the thickness inhomogeneities in the interface roughened pre-oxidation annealed samples[16]. The excess direct tunneling currents in the dry annealed samples are analyzed via a first order least squares fit to the Log(I) versus V data in the range of 2 - 4V. The intercept I_o, of the fit, is a measure of the physical dimensions of the MOS capacitor and will offset the I(V) upward for thinner films or larger area samples. All of our devices have similar area

Al contacts. The samples were grown such that they all had approximately the same optical thickness (±3Å) as determined by single wavelength ellipsometry. Therefore, any differences in I_o are a result of localized thickness inhomogeneities or some other transport mechanism.

The results for a dry pre-oxidation anneal study of the direct tunneling current as a function of anneal time are shown in Fig. 2 for degenerately doped p-Si(100) substrates. If we assume that the excess current is a result of direct tunneling, then the results show that for a zero time anneal, the interface is not significantly degraded. This is expected since the wafer has not had enough time to increase to a temperature where the surface will be etched by SiO formation. However after only a 30s anneal, the substrate morphology is degraded and there is a significant degree of excess current. As the anneal time is increased, I_o does not change. For anneal times up to 45min (not shown) the interface degradation saturates were all of the damage was established in less than 30s. These results are similar to experiments done on <111> oriented Si, and for various doping levels. The effects of varying the pH of a buffered HF (BHF) dip to remove the native oxide was examined. There are reports in the literature that indicate differences in the degree of hydrogen termination at the surface as a function of BHF pH[17]. These results were similar to those above. The annealing results were the same in either N_2 or Ar.

An analysis of the dielectric breakdown field as a function of dry pre-oxidation anneal times resulted in similar conclusions about the evolution of the interface roughness. Further indications that the interface degradation is due to surface roughness generated during the pre-oxidation anneal is the dependence of the breakdown field on the concentration of H_2O in the anneal ambient. Since the etching process of SiO formation was a result of a low

oxidant concentration, we attempted to study the effects on the roughening mechanism by varying the oxidant concentrations in the anneal. Figure 3 shows a study of the dependence of breakdown field on the concentration of H_2O as a controllable oxidant added to the anneal ambient in the pre-oxidation anneal ambient. Although the scatter in the data is large, there is some indication that the breakdown field does depend on the H_2O content of the anneal gas in the low concentration region. The dependence saturates as the water level is raised further.

It is possible that the addition of water to the anneal ambient can allay the etching process, resulting in a smoother interface and higher breakdown field. If this hypothesis is correct, we should be able to monitor the interface roughness as a function pre-oxidation anneal conditions. To this end we use the amplitudes of the quantum oscillations in Fowler-Nordheim tunneling currents as a probe of the interface roughness in ultrathin MOS tunneling capacitors.

II. EXPERIMENTAL

MOS capacitor structures were grown by thermal oxidation of Si in a double walled tube furnace. All of the silicon substrates were degeneratively doped p-Si(100) and were RCA cleaned[18] prior to processing. The native oxides were removed using a buffered hydrofluoric acid solution of pH~9. All pre-oxidation anneal times were held constant at 120s and at a furnace temperature of 800°C. Water concentrations in the anneal ambient was delivered through a bubbler and measured with a hygrometer at the tube furnace gas exit. All oxidations were carried out in "dry" O_2. The oxidant was turned on simultaneously

with turning off the wet anneal. It took about 5min for the residual water concentration of the anneal gas to reduce to levels of <10ppmV for all initial concentrations of H_2O in the oxidizing ambient. The average oxidation time at 800°C was 30min, such that we grew a constant oxide film thickness of 50±3Å, as measured by ellipsometry (N=1.7+0i)[19]. Aluminum injecting contacts of area $1 \cdot 10^{-3} cm^2$ were thermally evaporated through a shadow mask onto the oxide. Back contact was made via a GaIn eutectic paste. There was not any post metalization anneal.

The tunneling currents were measured using a ramped voltage technique. The substrate was biased into accumulation using a programmable bipolar operational amplifier power supply. The ramp rates were held constant at about 0.3V/s. Tunneling currents were sensed at the gate contact with a logarithmic picoammeter. Data collection and analysis was automated through a personal computer.

III. RESULTS AND DISCUSSION

A. Fowler-Nordheim Tunneling

For a sample that does not receive a dry pre-oxidation anneal, illustrated in Fig. 1 (closed square), the device exhibited neither excessive direct tunneling currents nor premature dielectric breakdown. As the field across the device is increased such that the separation of the Fermi levels in the contacts is greater than the oxide-anode work function, the electrons will tunnel through most of the barrier and into an energy level within the conduction band of the film. This is called FN tunneling[20], with the corresponding I(V) characteristic:

$$I(V) = A\left(\frac{V_{ox}}{d_{ox}}\right)^2 \exp\left[-\frac{Cd_{ox}}{V_{ox}}\right] \quad . \tag{3}$$

The constants A and C are dependent on the exact model used for the determination of T(E) but are typically related to the electron effective mass in the insulator and the injecting contact work function potential. Because of the trapezoidal shape of the barrier, the tunneling distance, d_T, will depend on the electric field across the insulator as:

$$d_T = \frac{d_{ox}\phi_m}{qV_{ox} + \Delta\phi_{ms}} \quad , \tag{4}$$

where ϕ_m is the injecting metal contact work function energy and $\Delta\phi_{ms}$ is the metal-semiconductor work function difference and V_{ox} is the voltage drop across the insulator of thickness d_{ox}.

To analyze the I(V) curve for a FN tunneling mechanism we need to rearrange Eqn. (4) into a linear relationship:

$$Ln(J_{FN}/F_{ox}^2) = Ln(A) - C/F_{ox} \quad . \tag{5}$$

The constants that describe the slope, C, and intercept, A, can be obtained by a least squares analysis of Eqn. 6. If we plot the data from an I(V) characterization using the FN coordinates $1/F_{ox}$ vs $\ln(J/F^2_{ox})$ we get the curve plotted in Fig 4. The solid line is the data and the dashed line the linear fit used to determine C and A. For this sample we evaluated

a slope of C = 2.7V/Å and intercept A = $1.35 \cdot 10^{-5}$. It is typical that the value of A varies a great deal from sample to sample and this is partially due to the difficulty in fitting the FN data to a straight line for ultrathin samples, as will be discussed below.

The electric field across the oxide, F_{ox}, must also be determined. For thick films where the FN conduction fields require very large applied voltages, V_{app}, the oxide field is assumed to be the given by $F_{ox} = V_{app}/d_{ox}$. But for thin films, like the samples studied here, the FN currents can be measured at the FN tunneling onset of V_{app}=4.42V ($V_{ox}=\phi_b$). For these smaller voltages, the errors induced by neglecting the work function differences of the metal and semiconductor contacts, W_{ms}, and the band bending effects at the semiconductor interface, become significant. At an applied bias of 5V across a 50Å n^+-poly/SiO_2/p^+-Si(100) sample, the voltage drop across the oxide is only ~3.5V. To analyze the FN data or calculate the dielectric breakdown field, one needs to convert the applied fields to the oxide fields. The extraction of F_{ox} from V_{app} is obtained using an iterative algorithm[21] to solve:

$$F_{ox} = \left(V_{app} + \phi_m - \phi_{so} - \alpha F_{ox}^{2/3}\right)/d_{ox} \quad , \tag{6}$$

where the term containing α describes the field dependence of the Si band bending[22].

B. Quantum Oscillations

Upon tunneling a distance, d_T, through the insulating barrier, electrons are injected into the oxide conduction band and ballistically propagate toward the anode. The electron

wave packets can reflect from the abrupt change in potential at the Si/SiO_2 interface. As long as there are no phase randomizing collisions within the propagation cavity of he MOS structure, the incident and reflected electron wave will interfere. For insulating film thicknesses that are on the same order of magnitude of the electrons inelastic mean free path, l_r, quantum oscillations have been observed in FN tunneling currents[23]. From Fig. 4 we can see the oscillations superimposed on the linear representation of our I(V) data. To elucidate the characteristics of the oscillations, we must subtract off the background FN current. Once the FN current density, J_{FN}, is calculated, it is divided into the data set, i.e. $\ln(J_{exp}) - \ln(J_{FN}) = \ln(J_{exp}/J_{FN})$. The normalized current density is plotted versus the applied voltage in Fig 5. This representation of the data accentuates the oscillations and any deviations to the FN transport mechanism. The oscillations can be described by position, V_i, period and amplitude Θ_{ij}, where the subscripts i = 1,2...n indicate the different extremum, starting at i=1 for the minima just after the onset of FN tunneling. The positions of the oscillations will vary depending on the shape of the potential barrier in the structure. The amplitudes of the oscillations are also sensitive to the shape of the barrier and to processes that can reduce the interference term such as inelastic collisions.

The quantum oscillations have been studied as a function of water concentration in a pre-oxidation anneal followed by the thermal oxidation of the Si substrate. Figure 6 illustrates typical oscillation results where the concentration of H_2O in the N_2 pre-oxidation anneal ambient was varied. The anneal was for 120s followed by turning off the H_2O and turning on the O_2 to grow a 50Å oxide film. These results show that as the concentration of H_2O in the anneal ambient decreases, so does the amplitude of the low field oscillation.

It is tempting to ascribe the decrease in oscillation amplitude to a real decrease in the interference term in the MOS quantum well. It is proposed that the decrease in the electron interference is due to roughness at the anode interface. However, before we can interpret the shape of the oscillations we need to further investigate our extraction of the oscillations from the FN data. It is important to point out, that the period of the oscillations for these samples is longer than would be theoretically expected for a simple trapezoidal barrier. From the measured voltage difference between the first and second minima, we calculate an oxide thickness of ~42Å. However, the thickness we extract by interpreting the displacement currents is 55Å and thus larger than the optical thickness, of 51Å. If the difference in the oscillation period was due to a thinner effective thickness from the interface roughness, the displacement currents should also yield a thinner oxide film. We do not have an explanation for this observation. The separation of the extrema of the oscillations is not correlated to the concentration of H_2O in the anneal ambient.

The low field oscillations can be misinterpreted if there are additional transport mechanisms that have non trivial contributions to the total current. Because the currents in the low field FN region are so small, an error in subtracting off the displacement currents can be interpreted as altering the shape of the oscillations. Also, if there is a significant contribution to the FN current by a direct tunneling mechanism, the low field current will appear higher. Since we are proposing that interface roughness will yield an excess tunneling current contribution, especially for low fields, then we must examine this effect as a possible explanation of the oscillation results.

To test the possibility that additional and unaccounted for transport mechanisms are

responsible for the change in the low field oscillation amplitude we simulate the effect of excess current on the measured I(V) characteristic. As the excess current is increased we found that the amplitude of the first minima in the oscillations decreases and the position of the minima of the first oscillation moves to higher applied biases. The simulated addition of the excess current simply raises the level of the low field current, in effect washing out the observed oscillations. In Fig. 6 we show that the amplitude of the low field oscillation decreases in a similar way as the excess current model would predict. However, for most of the oscillation data we do not see the continuous shift of the minima's position as the H_2O concentration is decreased. In fact the positions of the extrema are fairly insensitive to the processing conditions. If the decrease in oscillation amplitude was due to excess current washing out the oscillations, then there should be a corresponding shift in the oscillation position, which there is not. Therefore, we believe that the decreases in the oscillation amplitude is not due to the washing out of the oscillations from an excess current mechanism.

It is interesting to note, that the position of the first minima for the sample that received the 10ppmV H_2O anneal (Fig. 6) is shifted slightly to more positive biases . This amount of water in the "wet" anneal is only slightly higher than the amount present during the "dry" pre-oxidation anneal studies. For the dry annealed samples there was a large and dominant excess direct tunneling contribution to the I(V). The data for the 10ppmV anneal most likely represents a convolution of the two oscillation damping mechanisms. If this phenomena is due to interface roughness, then this data can be interpreted as the transition of the surface morphology being composed of large scale thickness inhomogeneities, to a

morphology of smaller scale coherence destroying defects.

As mentioned above, the existence of the oscillations makes the accurate determination of the FN constants difficult. To accomplish the least squares analysis of the data in the FN coordinates we need to judiciously select the range of the data used in the fitting routine. For example we would certainly not include a region of the data set that is obviously displaying a transport mechanism other than that of FN tunneling. The oscillations will create a problem in choosing the positions for the beginning and end of the data set, even if they were symmetric about the background current. If the sample is such that we were only able to measure two extrema of the oscillations then there is no way to independently determine the FN fit. When there are more than two oscillation peaks, we have a chance at interpreting the FN currents.

The oscillation amplitudes are being used in the context of a comparative study of the effect upon a MOS capacitor of varying the water concentration in the pre-oxidation anneal. Therefore, as long as we are consistent in our interpretation of the FN constants and that there is not significant correlation of the independent variable with the choice of the fitting constants, we can use the shape of the oscillations to probe the differences in the samples. For consistency in the data interpretation we assumed that the 2nd, 3rd and 4th extrema of the oscillations were symmetric about the FN current. Although only 3 oscillation extrema are needed to guarantee the possibility of fitting the FN constants, we did not include the analysis of a device unless we were able to measure at least 4 extrema. By setting C and A so that this condition was met, the fitting constants varied from device to device on the same sample, and also from sample to sample. We typically average the

analyzed parameters over 6 to 10 devices per sample and the oscillations that are shown in Fig. 6 are taken from the data set that most accurately represents the survey of devices across the sample.

Figure 7 illustrates the dependence of the average oscillation amplitude for the first two extrema, Θ_{12}, on the concentration of H_2O in the N_2 pre-oxidation anneal. Most of the data is for a 120s anneal. The error bars indicate the standard deviation of scatter between devices on the same sample. It is shown that Θ_{12} is linearly dependent on the H_2O concentration up to about 60ppmV. The effect saturates for higher levels of water in the anneal ambient. If we assume that the oscillation amplitude is changing because we are changing the interface roughness with the additional water, then the saturation of the effect indicates that either the interface roughness is approaching a minimum or that the oscillations are no longer sensitive to changes in the interface morphology. When the tunneling transmission coefficient for the 50Å MOS capacitor is calculated, the system is modeled with an abrupt and infinitely smooth interface without any inelastic scattering of the electrons throughout the entire device. Although the calculated oscillation amplitudes depend on several adjustable parameters, the maximum amplitude for Θ_{12} they we have ever calculated is about 0.9. Therefore, whatever the exact reason for the apparent saturation of the effect illustrated in figure 7, it is likely that the oscillations will be insensitive to the changes since the interference term is near its maximum theoretical value. When we do extract a Θ_{12} from the data that is greater than 1.0 the slope of the FN constants is typically much lower than the average value. The difficulty in determining C and A result in the anomalously large oscillation amplitudes and the increased scatter in the data.

As mentioned above the interpretation of interface roughness is only valid if the arbitrary choice of fitting constants is not correlated with the independent variable or the analyzed parameters. The correlation matrix comparing the concentration of water in the anneal ambient (below 60ppmV), the slope of the FN fit C, and the oscillation amplitude θ_{12} is shown in Table I. The matrix elements C_{ij} are normalized to vary over the range ±1 where a value of zero indicates that there is a random relationship between the parameters represented by the subscripts. Below 60ppmV of water, the oscillation amplitude seems very well correlated with the level of H_2O in the anneal ambient. Also, the correlation between the FN slope C, and θ_{12} is not as significant. This gives us more confidence in our ability to objectively extract the quantum oscillations from the FN I(V) and that we can quantitatively compare the oscillations between samples receiving various levels of H_2O in the pre-oxidation anneal.

C. AFM Imaging

Our results indicate the hypothesis that a dry pre-oxidation anneal will result in a rough substrate surface, and that the addition of water to the anneal ambient will allay the roughening process. In order to directly test this hypothesis, we imaged the samples with atomic force microscopy (AFM). For the samples receiving a 30, 50 and 315ppmV H_2O in N_2 pre-oxidation anneal, the AFM samples were prepared along with the samples used for the oscillations measurements. To study the morphology of the oxide/silicon interface, the oxide film was removed with a quick etch in concentrated HF. Care was taken to ensure that the oxide removal procedure did not alter the Si surface[24]. The sample was dried and

loaded into the AFM immediately for analysis. The images obtained from an AFM are both sample and tip dependent, and therefore the same tip was used to image all three of the substrates. Imaging a Si surface in air with AFM is difficult, and it is rare when the entire surface is homogeneous. Often the quantitative analysis of the surface morphology can be subjective. Therefore, all scanning parameters were also held constant for each of the samples, such that any differences could be attributed to the surface structure. To minimize personal biasing the samples were coded so that the microscopist did not know the processing involved in preparing the samples. Figure 8(a - c) is the AFM images of the three samples mentioned above. All the samples were low pass filtered and 1st order fit to a plane to remove sample tilt. Upon visual inspection, it is shown that the roughness of the surface decreases as the concentration of water in the anneal increases. Although it is apparent that the 315ppmV image is much smoother than the other two it is difficult to discern the 30ppmV sample from the 50ppmV. To further avoid the subjective analysis of the microscopist, the Digital Instruments Nanoscope III roughness analysis software was used to interpret the morphologies. By surveying the sample with several images we observed that the surface was not homogeneous and that a comparison of roughness analyses for some of the images found that the 30ppmV sample was smoother than the 50ppmV. However, we used the average and ranges of the roughness parameters to draw our conclusions about the degree of roughness on the surface. Table II lists some of the typical results from the roughness analyses. The "z-range" is the difference between the maximum and minimum height recorded in the image. The "RMS-rough" is the root mean square of the roughness relative to a plane drawn through the surface at the mean surface height. The "%ΔArea"

is the percent difference of the integrated area of the image compared to the area of a square which would represent a perfectly flat surface (viz. $16nm^2$). By analyzing several line scans across the image we can measure the typical bump height from maximum to nearest minimum and the bump width between minima. Scanning lines are visible in the AFM images and are an artifact of the data acquisition. It is important that we analyze the line scans across the sample in the same direction for all the samples. We typically pick a direction parallel with the scanning lines so that they do not contribute to the line morphology. The dimensions of the bumps in the surface roughness are small but they are still about an order of magnitude larger than the vertical and horizontal resolutions of the AFM. The frequency spectrum of the line scans do not exhibit any peaks at a well defined coherence lengths. Rather they are mostly DC with a continuous tail indicative of a pseudo-random distribution of roughness. From these data, it was determined by the microscopist, without knowledge of the sample's identity that the order for increasing roughness R^{AFM}, as a function of water concentration was $R^{AFM}(315ppmV) < R^{AFM}(50ppmV) < R^{AFM}(30ppmV)$. A more detailed microscopy study of the wet pre-oxidation anneal is warranted, but we are confident that the addition of water to the anneal does decrease the relative roughness at the Si/SiO_2 interface.

D. Roughness and Scattering

Roughness at the anode interface of a MOS capacitor can alter the quantum oscillations by several different mechanisms. Inhomogeneities in the oxide film thickness will reduce the amplitude of the oscillations. Different film thicknesses will result in oscillations

with different periods and the oscillation extrema will shift. A sum over periodic functions of varying frequencies results in the amplitude of the resultant interference term decreasing as the spread of the frequency of the interfering waves increases. The amplitudes of all of the oscillations should decrease monotonically with increasing thickness inhomogeneities. Sune[25] et. al. models roughness at the Si/SiO_2 interface as thickness inhomogeneities of the oxide film. In their model the device is broken up into a finite number of equal area columns. The height of each column is described by a Poisson distribution of h steps, of height a, such that the local thickness is $d_l = h \cdot a$. The step size determines the degree of roughness where $a = 0$ would yield a flat interface and uniform oxide thickness. In order to calculate the shape of the quantum oscillations as a function of roughness, they determined the I(V) characteristic for each column of thickness d_l. Then the total current through the device is the summation over all of the columns with local current through d_l, I(V;a,h), modulated by the probability density for that column. This model not only predicts that the amplitude of the oscillations will decrease as the degree of roughness increases, but also the positions of the oscillations will shift. The distribution of column heights produces an oxide film with a thinner effective thickness than would be measured on a flat film. As the effective thickness decreases, the oscillations shift toward higher applied biases. This model could explain why the oscillations we observed in Fig. 6 have longer than expected periods and are shifted as if the film was thinner.

However, Sune's model predicts that all of the oscillations will decrease in amplitude as the thickness inhomogeneities increase. We do not observe a monotonic change of all the oscillation amplitudes when we vary the water concentration in the pre-oxidation anneal

ambient. Figure 9 shows the dependence of the oscillations amplitudes Θ_{12} (circle), Θ_{23} (square) and Θ_{34} (diamond) on the H_2O concentration in the anneal. As shown above, the Θ_{12} is linearly dependent on the water concentration in the pre-saturation regime. The lines represent a least squares linear analysis of the data. The characteristics of the higher field oscillations are not similar to the low field oscillation. The high field oscillation amplitudes are nearly insensitive to the water levels, and in fact seem to decrease as the H_2O concentration increases. This could not happen if the interface roughness was described simply by oxide thickness inhomogeneities.

As mentioned above, electron scattering in the bulk of the oxide or at the reflecting interface can destroy the coherence of the wave and the interference term will decrease. A decrease in the interference amplitude will only decrease the quantum oscillation amplitudes. It will not shift the positions of the oscillations or change their period. This is what we observe in the pre-oxidation anneal study. Although, there is some variations in the shape of the oscillations from device to device and sample to sample, there is no significant shifting of the positions of the oscillation extrema. This also implies that the effective thickness of the samples was fairly constant as would be necessary for a direct comparison of the oscillations.

For the electron wave packet to interfere in the MOS quantum well, it must maintain phase coherence throughout a distance traversing a round trip in the oxide conduction band. If the electron undergoes an inelastic phase randomizing collision within the cavity or at the interface, it will not add constructively to the total interference amplitude. If the electron elastically collides within the structure such that the resulting momentum vector has

significantly deviated from a parallel direction of the incident wave vector, then the electron will not effectively interfere. The interference of incident and reflected electron density depends on the overlap of the wave function before and after the collision.

The interactions of electrons with surface roughness has been studied by examining the resistivity of thin films as a function of film thickness[26]. It was shown that the resistivity of the film increased as the film thickness decreased. It was concluded that roughness at the surface of the film dominated the resistivity as the dimensions of the bulk became comparable to the thickness of the roughness layer. It has been shown theoretically[27] that scattering of the conduction electrons in metallic films with the surface roughness is responsible for the increase in resistivity. The intensity of electron scattering from the surface was shown to be dependent on the correlation distance describing the roughness. We can draw an analogy from the optical scattering of electromagnetic radiation with particles at a surface[28], with an electron wave packet scattering from surface roughness. The intensity of the scattered wave is dependent on the relative dimensions of the roughness compared to the wavelength of the incident wave. In the quantum oscillation analysis of the interface roughness of the MOS device, there is both a distribution of the dimensions of the roughness and a variation of the electron wave packet's wave length. The wavelength of the wave packet depends on the applied bias across the structure and are listed in Table III for three consecutive extrema in a typical pre-oxidation anneal I(V). In the absence of scattering in the bulk of the oxide the scattering intensity of the roughness as a function of the electron's wave length may be used to probe the degree and size distribution of the interface roughness. We are not aware of a suitable theoretical treatment

of this situation and thus can not quantify the analysis at this time. This model is a possible explanation for the results of Fig. 9 where the long wavelength electrons are more sensitive to the type of interface roughness we are measuring and the scattering intensity of the shorter wave length electrons is insensitive to the roughness.

We neither have a quantitative analysis for the electron-roughness interaction, nor are we sure if there is an experimental dependence of θ_{23} and θ_{34} on the apparent roughness measured by θ_{12}. If anything, it seems that the high field oscillation amplitudes are decreasing as the concentration of water in the anneal increases. For the samples exposed to a 10 or 20ppmV H_2O anneal, the high field oscillation amplitudes are about what they would be for a sample that was pushed into the oxidation furnace while dry O_2 was the ambient. As the water concentration is increased to greater than 30ppmV the oscillation amplitudes decrease to about half of their original value. We have previously established that the concentration of water in the oxidizing ambient drops off quickly after the water is turned off and the O_2 is turned on. However, it may be possible that there was a non-negligible concentration of water in the oxidizing ambient for the higher H_2O concentration pre-oxidation anneals. From the H_2O versus time profile we have confidence that the interfacial region of the oxide film was grown under the same oxidizing ambient no matter what the initial concentration of water. However, the bulk region of the films may have been grown with a varying degree of water in the oxidizing atmosphere.

Wet oxidations are responsible for the creation of water related traps in the oxide[29]. These traps can act as scattering centers in the bulk of the film. Scattering while the electron is tunneling through the oxide forbidden band should not influence the interference

of the electrons as they are propagating in the oxide conduction band. However, if the traps or impurities[30] were located with sufficient concentration in the propagation region, then they could be responsible for the electron scattering and the reduction of the interference amplitude. Table III lists the calculated propagation distances of the electron in the oxide conduction band for several biases in the I(V). The biases correspond to the positions of the extrema in the oscillations. The propagation distance for the low field oscillation is almost half that for the high field oscillations. Also, the propagation region for the low field oscillation is about 10Å closer to the interface. The concentration of water in the oxidizing ambient was most likely less than that present during the growth of the bulk of the film. It is proposed that electron scattering in the initial propagation region of the MOS structure under high bias dominates the scattering processes and that the quantum oscillations are insensitive to the nature of the interface.

IV. SUMMARY

We have shown that, although not without difficulty, the quantum oscillations can be extracted from the FN background current, and that the amplitude of the oscillations depended on the concentration of H_2O in the ambient gas of a pre-oxidation anneal followed by oxide growth to form MOS capacitors. It was proposed that changes in the oscillation amplitudes as a function of water concentration was a result of scattering in the bulk of the film and at the anode interface. The scattering at the interface is most likely the result of interface roughness. The shape of the oscillations result from a convolution of the various scattering mechanisms and the scattering intensity as a function of the interaction

of the incident wave packet with the scattering center. This method has potential as a non-destructive probe of interface roughness in MOS tunneling structures.

ACKNOWLEDGMENTS

This work was supported in part by the Office of Naval Research. We also appreciate partial funding for one of us (JCP) through a fellowship from the Semiconductor Research Corporation Education Alliance.

1. F.J. Grunthaner and P.J. Grunthaner, Materials Sciences Reports **1**, (North Holland, Amsterdam, pp. 65-68.

2. E. Kobeda and E.A. Irene, J. Vac. Sci. Technol. B **6**, 574 (1988).

3. S.R. Kasi, M. Liehr, P.A. Thirty, H. Dallaporta and M. Offenberg, Appl. Phys. Lett. **59**, 108 (1991).

4. M. Liehr, M. Offenburg, S.R. Kasi, G.W. Rubloff and K. Holloway, Extended Abstracts of the 22nd Conference on Solid State Devices and Materials, 1099 (1990).

5. T. Takahagi, I. Nagai, A. Ishitani, H. Kuroda and Y. Nagasawa, J. Appl. Phys. **64**, 3516 (1988).

6. S. Watanabe, N. Nakayama and T. Ito, Appl. Phys. Lett. **59**, 1458 (1991).

7. P. Gupta, V.L. Colvin and S.M. George, Phys. Rev. **37**, 8234 (1988).

8. F.W. Smith and G. Ghidini, J. Electrochem. Soc. **129**, 1300 (1982).

9. E.A. Irene, Phil. Mag. B **55**, 131 (1987).

10. M. Offenberg, M. Liehr, G.W. Rubloff and K. Holloway, Appl. Phys. Lett. **57**, 1254 (1990).

11. V.A. Yakovlev, Q. Liu and E.A. Irene, J. Vac. Sci. Technol. A **10**, 427 (1992).

12. C.K. Chow, J. Appl. Phys. **34**,2599 (1963).

13. E. Farres, J. Sune, I. Placencia, N. Barniol and X. Aymerich, Phys. Stat. Sol. A **113**, 83 (1989).

14. G. Binnig and H. Rohrer, IBM J. Res. Develop. **30**, 355 (1986).

15. R. Stratton, J. Phys. Chem. Solids **23**, 1177 (1962).

16. J.C. Poler and E.A. Irene, Chemical Surface Preparation, Passivation and Cleaning for Semiconductor Growth and Processing, edited by R.J. Nemanich, C.R. Helms, M. Hirose and G.W. Rubloff (Mater. Res. Soc. Proc. **259**, Pittsburgh, 1992) pp. 93-98.

17. G.S. Higashi, R.S. Becker, Y.J. Chabal and A.J. Becker, Appl. Phys. Lett. **58**, 1656 (1991).

18. W. Kern and D.A. Puotinen, RCA Rev. **31**, 187 (1970).

19. S. Chongsawangvirod, E.A. Irene, A. Kalnitsky, S.P. Tay and J.P. Ellul, J. Electrochem. Soc. **137**, 3536 (1990).

20. R.H. Fowler and L. Nordheim, Proc. Royal Soc. **A119**, 173 (1928).

21. J. Maserjian, The Physics and Chemistry of SiO_2 and the Si-SiO_2 interface, edited by C.R. Helms and B.E. Deal, (Plenum, New York, 1988) p. 501.

22. J. Maserjian, G. Petersson and C. Svensson, Solid-State Electron. **17**, 335 (1974).

23. J. Maserjian, J. Vac. Sci. Technol. **11**, 996 (1974).

24. K.K. McKay, Private Communication.

25. J. Sune, I. Placencia, E. Farres, N. Barniol and X. Aymerich, Phys. Stat. Sol. (a) **109**, 479 (1988).

26. P. Wissmann and K. Muller, Surface Physics 77, Springer Tracts in Modern Physics, (Springer, New York, 1975), Chap. 1.

27. K.M. Leung, Physical Review B **30**, 647 (1984).

28. R. Schiffer, Scattering in Volumes and Surfaces, Edited by M. Nieto-Vesperinas and J.C. Dainty (North Holland, Amsterdam, 1990), pp. 223-38.

29. E.H. Nicolian, C.N. Berglund, P.F. Schmidt and J.M. Andrews, J. Appl. Phys. **42**, 5654 (1971).

30. N.W. Ashcroft and N.D. Mermin, Solid State Physics, (Saunders College, Philadelphia, 1976), pp. 320-22.

Figure 1: Typical I(V) characteristics for Al/SiO_2/Si devices. Differences in processing are displayed in the different tunneling mechanisms. Direct (●) and FN (■) tunneling are compared.

Figure 2: The y-intercept of a fit to the tunneling currents through MOS capacitors that received a pre-oxidation anneal in dry N_2, as a function of anneal time.

Figure 3: Average breakdown field for pre-oxidation annealed MOS capacitors as a function of H_2O concentration in the anneal ambient. Anneal time is 120sec and oxide thickness ~50Å.

Figure 4: Fowler-Nordheim plot of tunneling data (solid) compared to the least squares linear fit (dashed).

Figure 5: Quantum oscillations in the FN tunneling currents for a 50Å thick MOS capacitor versus applied voltage. Oscillation amplitude Θ_{ij}, position V_i, and period are illustrated.

Figure 6: Quantum oscillations in FN I(V) of ~50Å thick MOS capacitors on p^+-Si(100). Oscillation amplitude is dependent on the concentration of H_2O in the pre-oxidation anneal.

Figure 7: Average oscillation amplitude Θ_{12} vs H_2O concentration in pre-oxidation anneal ambient. Θ_{12} depends linearly on H_2O conc., then saturates for higher water levels.

Figure 8: Typical AFM images of wet pre-oxidation annealed samples after oxide removal. Images of 30ppmV (a), 50ppmV (b) and 315ppmV (c) anneal samples were low pass filtered and fit to a plane.

Figure 9: Oscillation amplitudes for Θ_{12} (●), Θ_{23} (■) and Θ_{34} (♦) as a function of water concentration in the pre-oxidation anneal ambient. Linear fits to the data below 60ppmV are shown.

Table I Correlation matrix of H_2O concentration, FN slope C, and oscillation amplitude θ_{12} for the data of the anneals that contain less than 60ppmV of H_2O.

C_{ij}	H_2O Conc.	C_{FN}	θ_{12}
H_2O Conc.	1	-0.317	0.969
C_{FN}	-0.317	1	-0.370
θ_{12}	0.969	-0.370	1

Table II Summary of roughness analysis results for pre-oxidation annealed Si surfaces.

ppmV	z-range Å	rms-rough. Å	%ΔArea	bump height Å	bump width Å
30	8.1	1.1	93.3	1 - 5	4 - 7
50	7.7	1.0	56.0	1 - 2	4 - 6
315	5.8	0.7	30.9	0.5 - 1	7 - 14

Table III Positions of the oscillations from a 50ppmV pre-oxidation annealed MOS capacitor V_{app}, the propagation distance in the oxide conduction band d_p, and the electron's wavelength at the interface λ.

Osc #	V_{app} (Volts)	d_p (Å)	λ (Å)
1	4.65	12.5	21
2	5.43	18.6	12.7
3	5.71	20.0	11.4

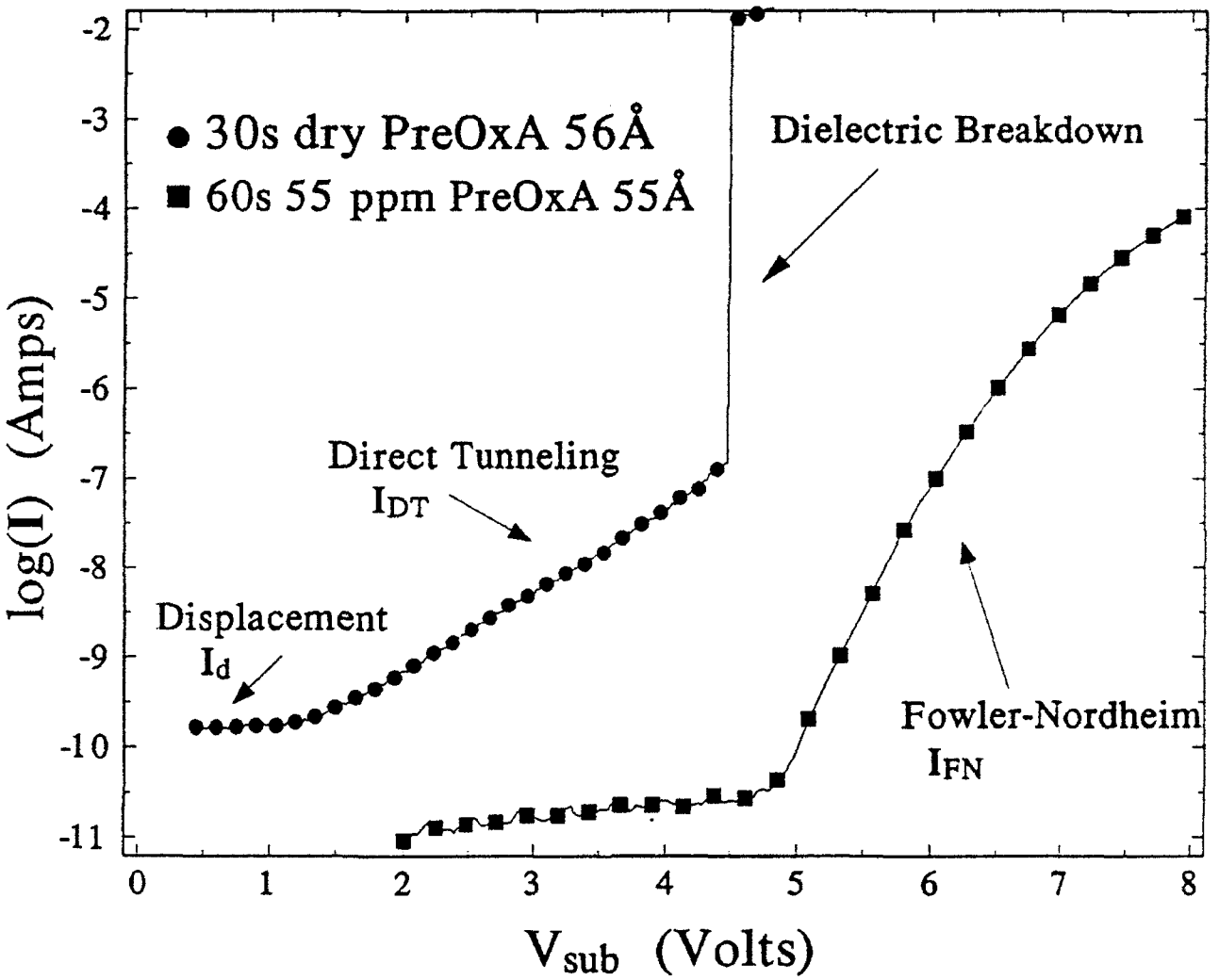

Fig. 1 Jordan Poler

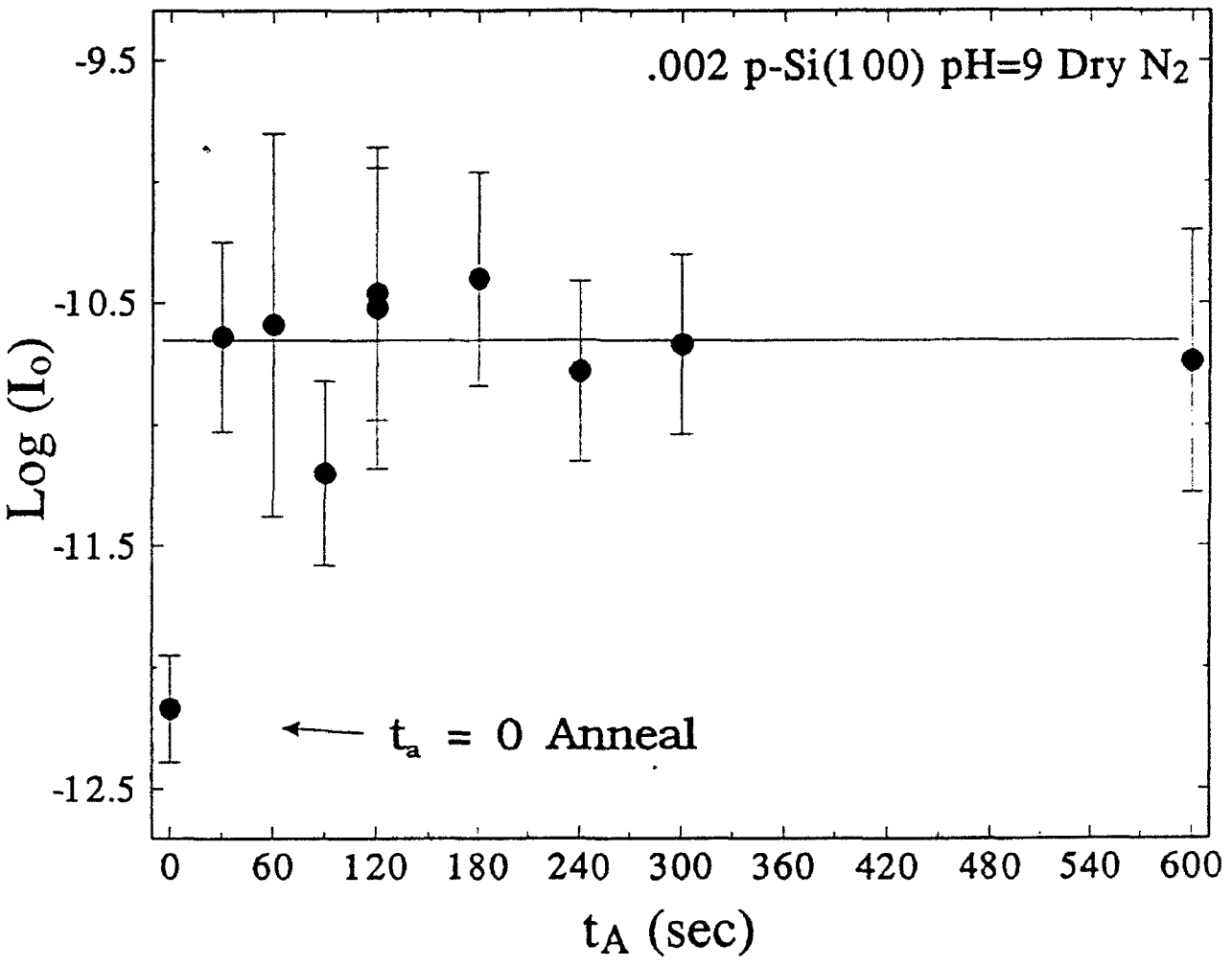

Fig. 2 Jordan Poler

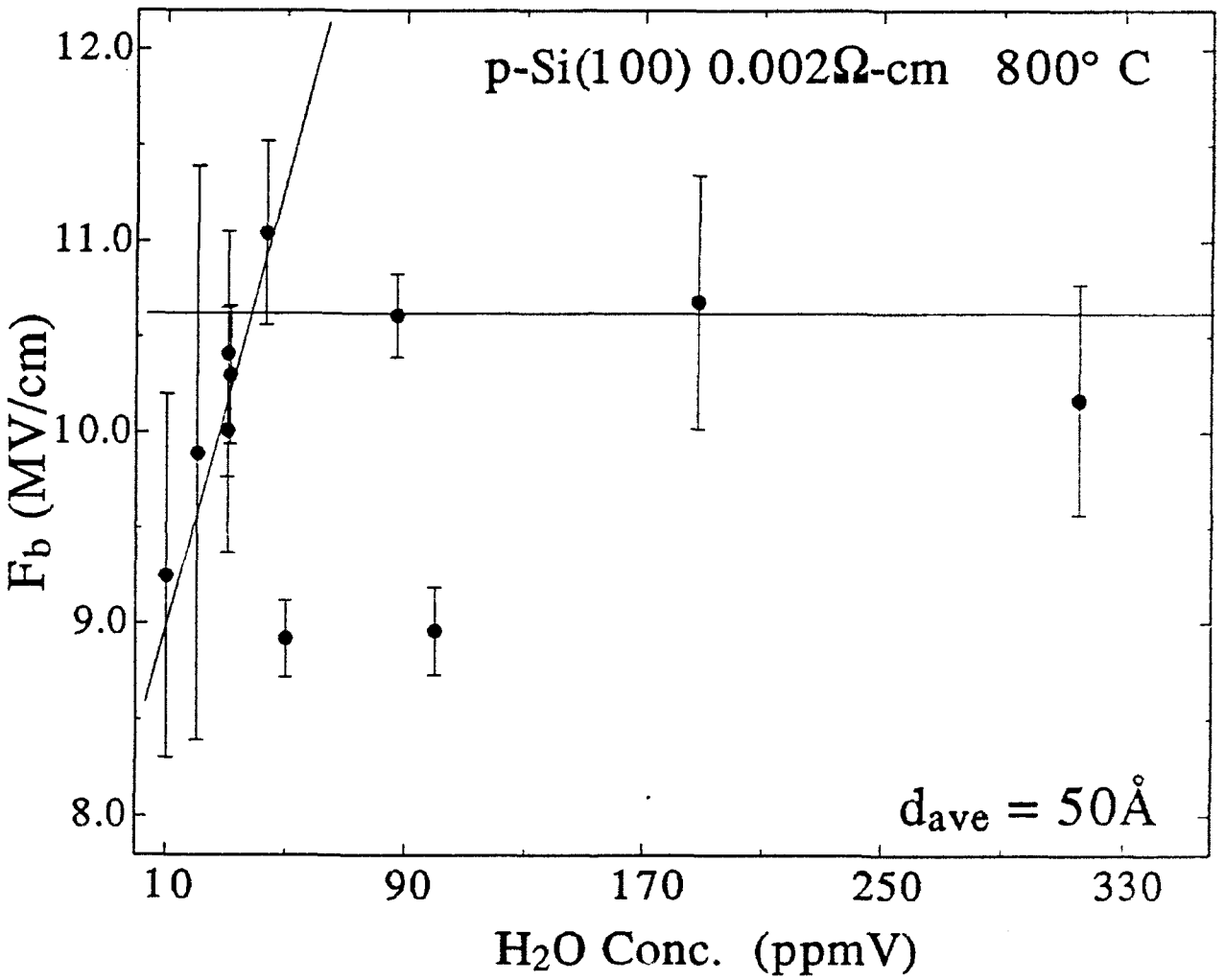

Fig. 3. Jordan Poler

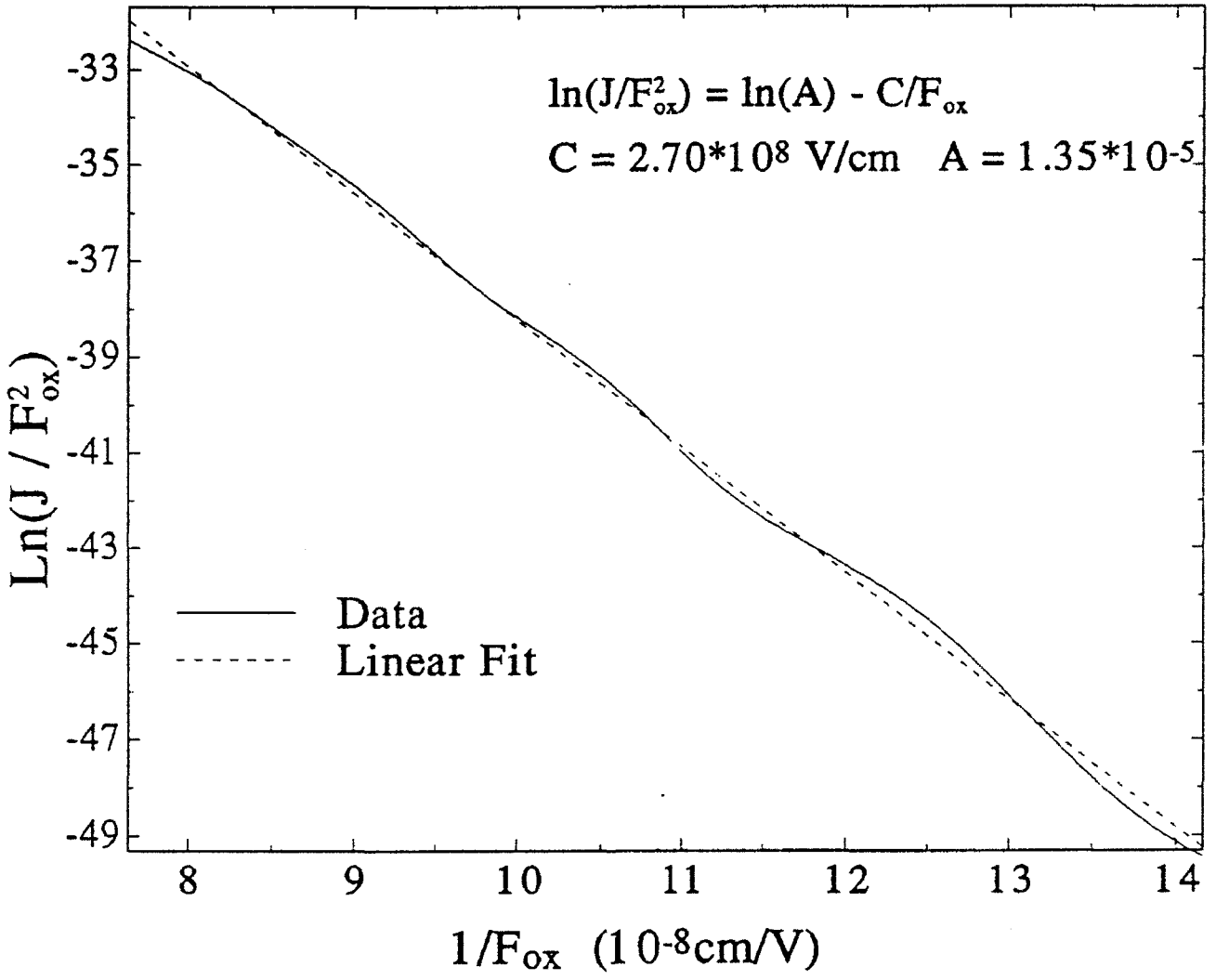

Fig 4. Jordan Poler

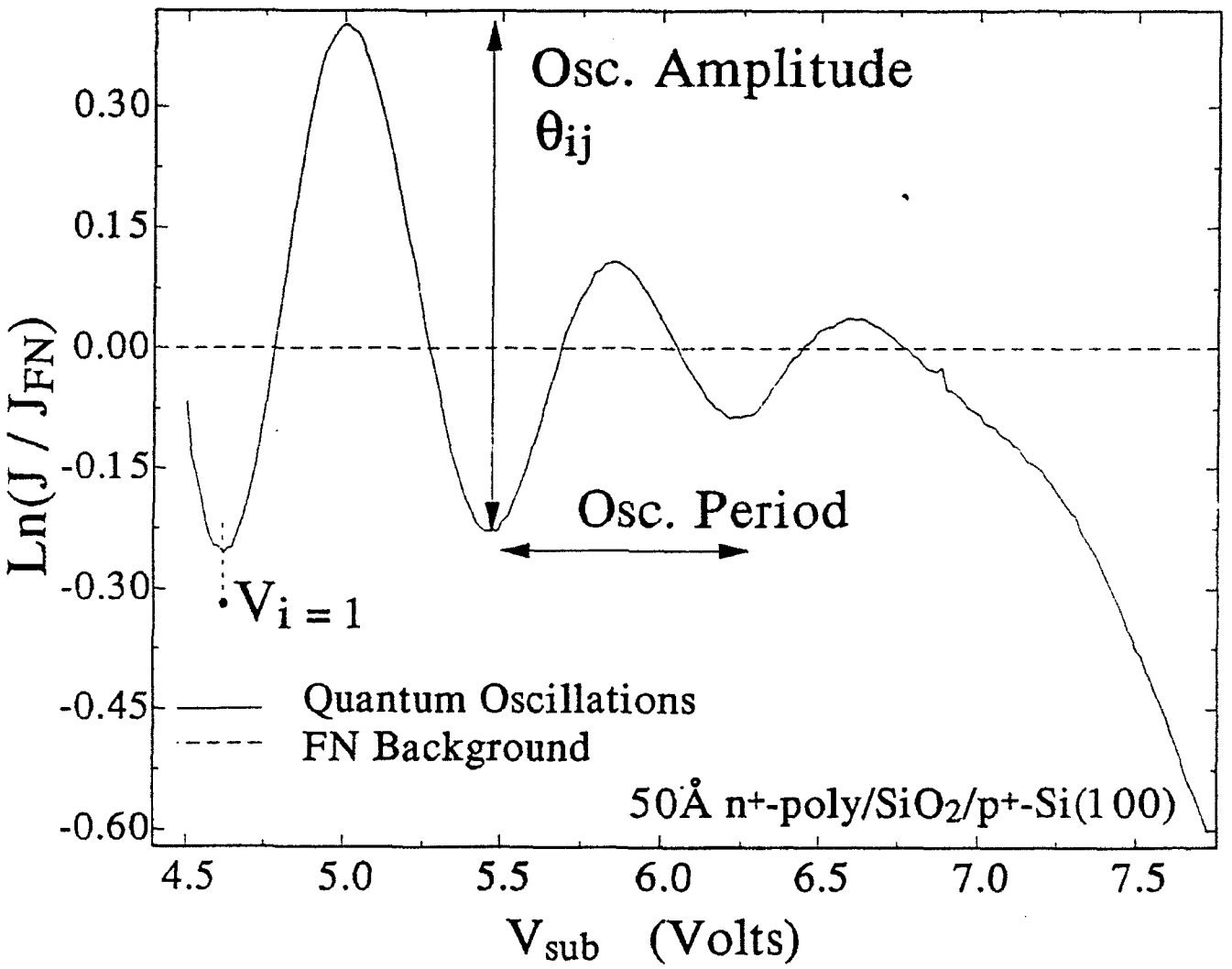

Fig. 5 Jordan Poler

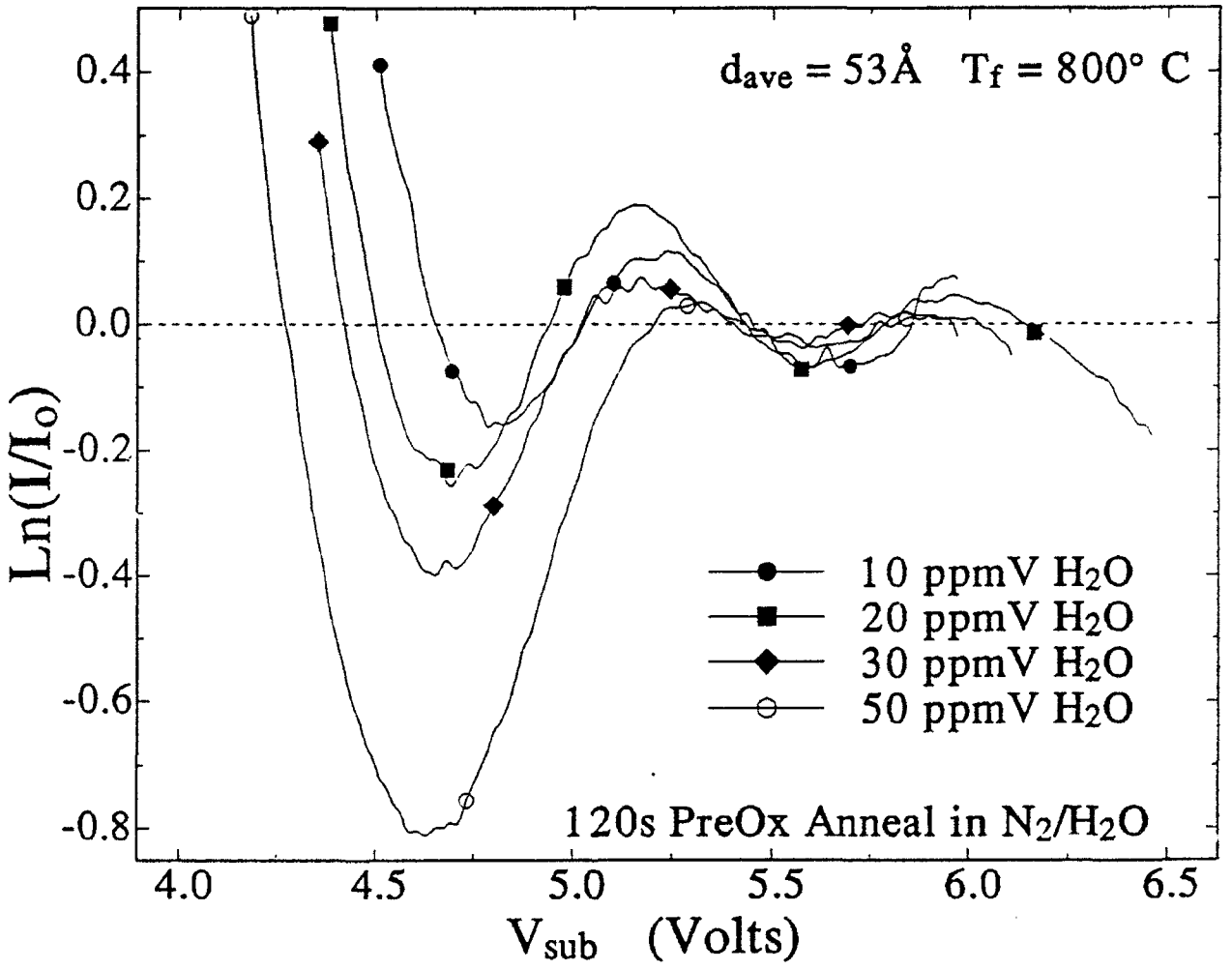

Fig 6 Jordan Paper

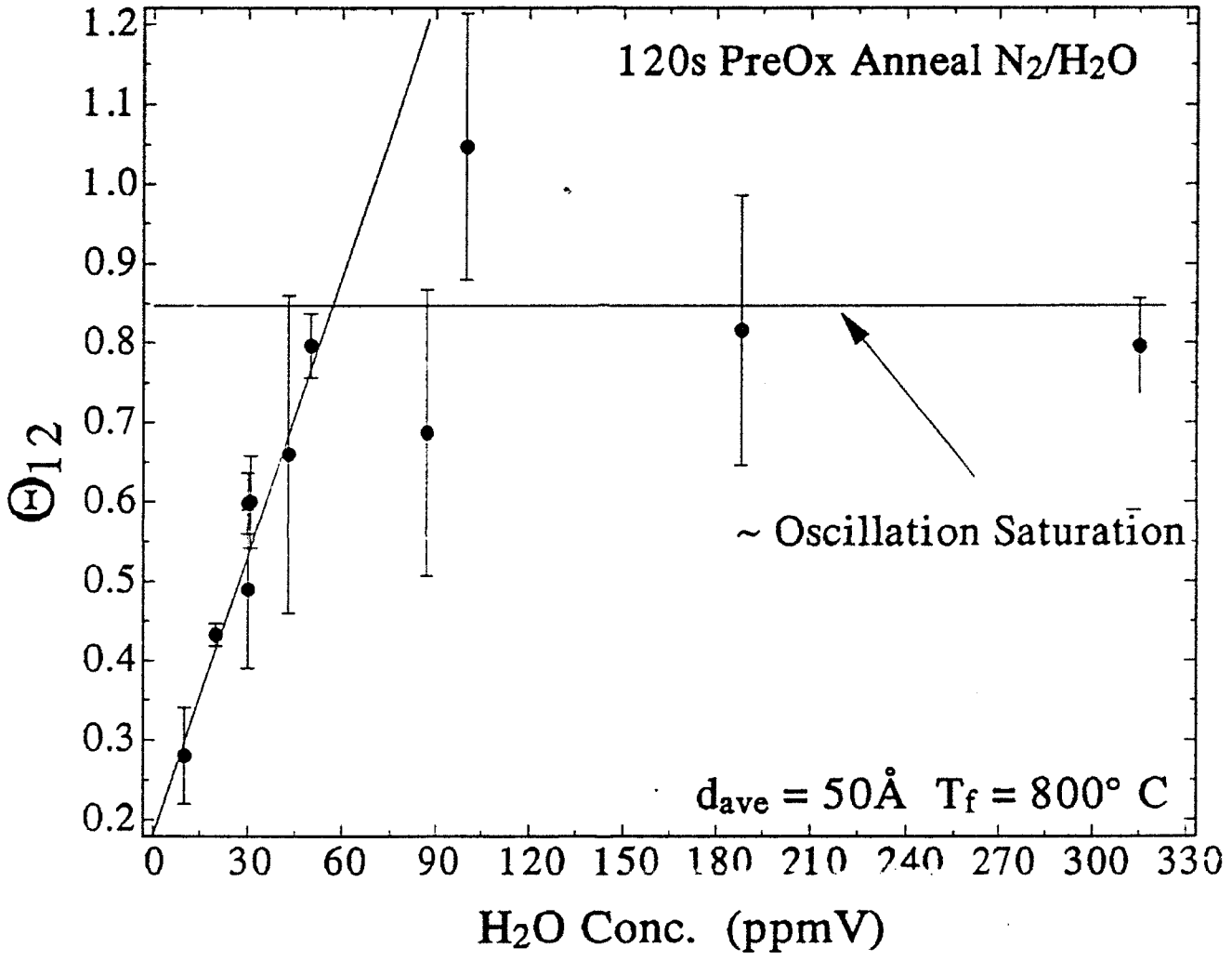
120s PreOx Anneal N_2/H_2O
~ Oscillation Saturation
d_{ave} = 50Å T_f = 800° C
Θ_{12}
H_2O Conc. (ppmV)
1.2
1.1
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0
30
60
90
120
150
180
210
240
270
300
330

Fig. 7 of 11 01 -

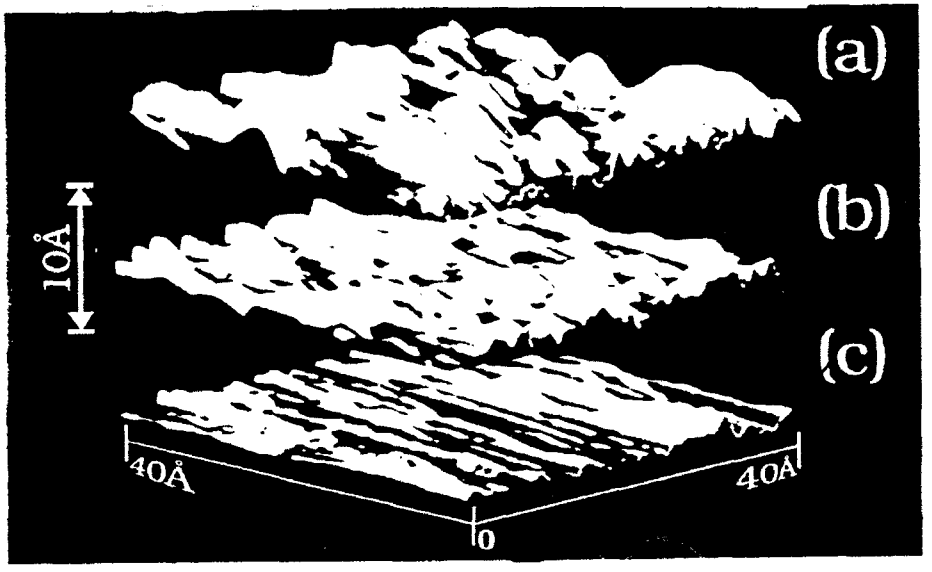

Fig 8. Jordan Poler

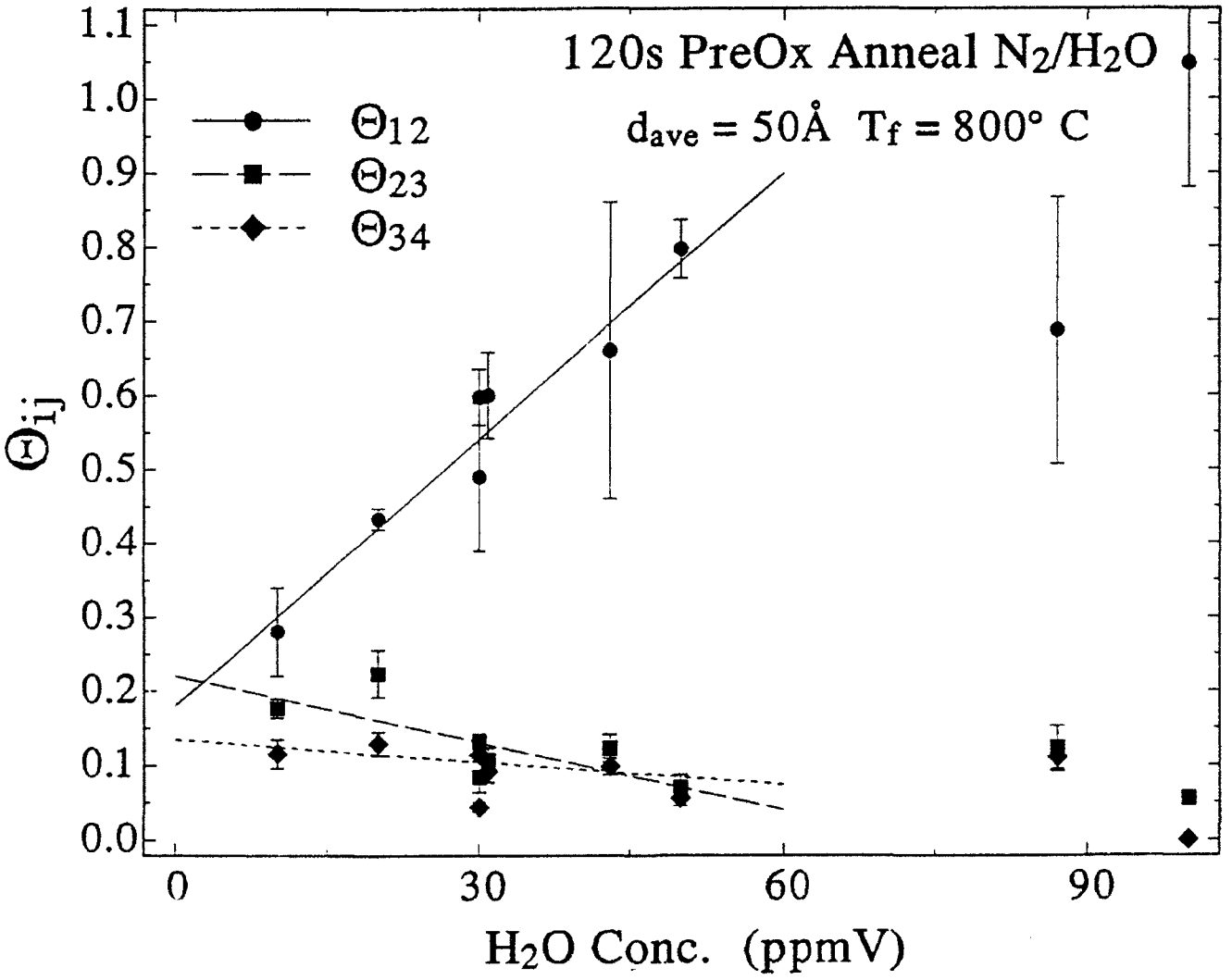

Fig 9 Tab Polar